

RFP: GLOBAL SEO TECHNICAL ROLLOUT

Project Overview: Rollout Technical SEO updates to sites worldwide on a region-by-region basis. This project will have to take into consideration locale/brand nuances to ensure minimal regressions.

Scope of Work: The scope of this project includes the implementation of 6 technical features to support SEO (listed below.) Due to the size/scope of this project, we will need a project manager to oversee and communicate all tasks to business owners.

- [Meta Data Availability & Function](#)
- [Rich Snippets](#)
- [No Index/ No follow](#)
- [Canonical URLs](#)
- [Robots.Txt File](#)
- [Sitemaps/HREFLANG](#)

Business Rules & Technical Requirements can be found by clicking each of the above features.

Business Rules & Technical Requirements:

Meta Data Exposure & Function: The goal of this feature is to ensure that the meta description and title tag fields are all exposed in .NET and the CMS and are working to result in updating site HTML, specifically <meta name="Description" content="value entered in .NET or CMS". The secondary goal of this feature is to insert backup logic to propagate these fields for SPPs using .NET short description should no copy exist for meta description.* Each brand/locale .NET needs to be checked to ensure field presence (exposing fields for the brands without presence.) From there, dummy copy needs to be entered-we cannot test with the copy that is defaulted in these fields. (Example: Use "This is test copy for a product".) To validate, we'll need to ensure that the copy entered shows on the site 24-48 hours later for each field. After this is verified, testing should delete these fields to ensure that the default logic takes action. Default logic should be implemented to pull in the product short description for all SPPs. Please reference full **SEO CMS dotNET mapping** guide for specific business rules and how to update per page type and brand.

Fields	Verified in
Meta Description	Site HTML specifically <meta name="Description" content="value entered in .NET or CMS".
Title tag	Title tab in the site browser

- [Example Ticket](#)
- Reference: [3.8 of Platform Req Document](#)
- *Default description is not available for the below brands because they do not have a short description field. Adding a new default field would require a new scope.
 - Glamglow
 - Bobbi Brown
 - Estee Lauder

Rich Snippets: The goal of this feature is to standardize what relevant snippets show in search results. QA validation results in a “Rich Snippet” (AKA “Structured Data”) code being implemented on all pages in the correct format. QA should include the verification of 10 different product types for each brand: Shaded, Non-shaded, sized, unsized). This feature should ONLY be added for sites with Ratings & Reviews. The schema pulls from data specific to the site and will reflect local language when used. For example: product name will only be translated if the site translates the product name.

- [Example Ticket](#)
- Reference: [3.7 of Platform Req Document](#)

No Index/No Follow: The goal is to ensure all onsite search results pages are blocked from offsite search results. This is achieved by adding this meta tag to the head of the search page. Upon completion, this fix should show up within a few hours. Check production robots.txt file to ensure updates are reflected by typing site URL/robots.txt (i.e.maccosmetics.com/robots.txt)

- [Example Tickets](#)
- Reference: [3.5 of Platform Req Document](#)

Canonical URLs: The purpose of this feature is to add a call to help alleviate the duplicate content caused by query parameters. Please add the call snippet to all sites on all products and upon completion (per brand), request content query is run in google search console by global team, note updates take 1-2 weeks to show up. HTML Call: <link rel = “canonical” href=<https://www.maccosmetics.com/product/13854/53464/products/makeup/lips/lipstick/retro-matte-liquid-lipcolour-metallics>>

- [Example Ticket](#)
- Reference: [3.1 of Platform Req Document](#)

Robots.txt File: The purpose of this feature is to ensure that anything that lives under the “/shared” directory is blocked from search engine crawling. Add “/shared” items to the brand specific robots.txt file and upon completion, the URLs should show in the production robots.txt file. Updates should also be validated by typing site URL/robots.txt (i.e. maccosmetics.com/robots.txt) and verifying presence of “/shared” URLs.

- [Example Ticket](#)
- Reference: [3.4 of Platform Req Document](#)

Sitemaps/HREFLANG: Validate that a sitemap is live by visiting the site’s extension of: /sitemap.xml (i.e.maccosmetics.com/sitemap.xml). Confirm that every map includes HREFLANG tags in the Language-Country format, and that sitemaps are in adherence with details from the relative material linked below.

- Reference: [3.10 of Platform Req Document](#)

3.1 Canonical URLs

A canonical URL allows a site to tell search engines that certain similar URLs are actually one and the same. Sometimes you have products or content that are accessible under multiple URLs. Using a canonical URL, these separate references can exist without harming your rankings as duplicate content might dilute your ranking ability. Many sites have encountered issues with multiple versions of the same content on slightly different URLs. This creates a number of problems:

1. Search engines don't know which version(s) to include/exclude.
2. Search engines don't know whether to direct the link metrics (trust, authority, anchor text, link juice, etc.) to one page, or keep it separated between multiple versions.
3. Search engines don't know which version(s) to rank for query results.

When this happens, sites suffer rankings and traffic losses. The SEO Task Force would like to ensure that any new platform provides us with a way to alleviate the duplicate content problem.

Canonical URL Case Study: DrupalGem & Duplicate Content

Duplicate content is an issue on DrupalGem. To address this issue, the UK team created a meta-tag function that is able to be applied to any DG site. This meta-tag function lives within the Global Codebase. In order to received the benefits of this function you will need to first turn them on for every site.

Example

1. On MAC UK we include the call to 'buildCanonicalURL' to the following file:
 - a. /u01/perlgem_www/configuration/Mac/GlobalServices_UK_MS_UK/prodcat_rpc.tpl
2. Firstly the call was added to the spp_product_fields array .
3. Within that file we also add a call to this sub routine at the end of the rated_product_fields array.
4. The function will return a meta-tag string that contains the URL without any parameters and configured to all lower case.

3.4 Robots.txt

Robots.txt is a text file webmasters create to instruct web robots (typically search engine robots) how to crawl pages on their website. These crawl instructions are specified by "disallowing" or "allowing" the behavior of certain (or all) user agents. We must exclude the proper URLs within the robots.txt files based on the platform URL structure.

General rules around website section exclusion apply to the new platform. We would like to continue to disallow the following sections:

Do not crawl these section (Disallow)

- Onsite search
- Checkout
- Account
- Site down maintenance page
- Any tools directories (example on DG platform /shared/*)

Case Study: DrupalGem & Robots.txt

All DrupalGem websites are setup with default settings in their robots.txt file. In theory, these settings should be reviewed with every site implementation to ensure they are relevant to the website and market.

```
*****
*****
```

```
User-agent: *
Crawl-delay: 10
# Directories
Disallow: /includes/
Disallow: /misc/
```

Disallow: /modules/
Disallow: /profiles/
Disallow: /scripts/
Disallow: /themes/
Files
Disallow: /CHANGELOG.txt
Disallow: /cron.php
Disallow: /INSTALL.mysql.txt
Disallow: /INSTALL.pgsql.txt
Disallow: /INSTALL.sqlite.txt
Disallow: /install.php
Disallow: /INSTALL.txt
Disallow: /LICENSE.txt
Disallow: /MAINTAINERS.txt
Disallow: /update.php
Disallow: /UPGRADE.txt
Disallow: /xmlrpc.php
Paths (clean URLs)
Disallow: /admin/
Disallow: /comment/reply/
Disallow: /filter/tips/
Disallow: /node/add/
Disallow: /search/
Disallow: /user/register/
Disallow: /user/password/
Disallow: /user/login/
Disallow: /user/logout/
Disallow: /account/
Disallow: /checkout/
Disallow: /templates/
Disallow: /includes/
Disallow: /product/15579/41818/checkout/Chubby-Contour-Kit/Chubby-Contour-Kit
Paths (no clean URLs)
Disallow: /?q=admin/
Disallow: /?q=comment/reply/
Disallow: /?q=filter/tips/
Disallow: /?q=node/add/
Disallow: /?q=search/
Disallow: /?q=user/password/
Disallow: /?q=user/register/
Disallow: /?q=user/login/
Disallow: /?q=user/logout/
Disallow: /shared/
Disallow: *.xml
Allow: /sitemap.xml

3.5 Noindex / Nofollow

You can prevent a page from appearing in search engines by including a noindex meta tag in the page's HTML code, or by returning a 'noindex' header in the HTTP request. Using noindex meta tag is useful if you don't have root access to your server, as it allows you to control access to your site on a page-by-page basis.

Nofollow is a value that can be assigned to the rel attribute of an HTML a element to instruct some search engines that the hyperlink should not influence the ranking of the link's target in the search engine's index.

All onsite search related pages should contain the noindex/nofollow meta tags.

3.7 Rich Snippets

Rich Snippets provide extra information that is shown on a search engine results page (SERP). These details add an extra layer of glamour to your search output. HTML markup can be seen from search results snippets: for example, H2 heading often appears in a rich snippet. How our HTML market up is presented might give us the edge we need when it comes to organic SEO.

Case Study: DrupalGem & Rich Snippets

The code to enable rich snippets is readily available on the DrupalGem platform. To get Rich Snippets working each site is required to add a Script Tag to all product pages. An example of this can be seen below:

Example:

```
1. <script type="application/ld+json">
2. {
3.   "@context": "http://schema.org/",
4.   "@type": "Product",
5.   "name": "Executive Anvil",
6.   "image": "http://www.example.com/anvil_executive.jpg",
7.   "description": "Sleeker than ACME's Classic Anvil, the Executive Anvil is perfect for the business
traveler looking for something to drop from a height.",
8.   "mpn": "925872",
9.   "brand": {
10.    "@type": "Thing",
11.    "name": "ACME"
12.  },
13.   "aggregateRating": {
14.    "@type": "AggregateRating",
15.    "ratingValue": "4.4",
16.    "reviewCount": "89"
17.  },
18.   "offers": {
19.    "@type": "Offer",
20.    "priceCurrency": "USD",
21.    "price": "119.99",
22.    "priceValidUntil": "2020-11-05",
23.    "itemCondition": "http://schema.org/UsedCondition",
```

```
24.  "availability": "http://schema.org/InStock",
25.  "seller": {
26.    "@type": "Organization",
27.    "name": "Executive Objects"
28.  }
29. }
30. }
31. </script>
```

Replace the content within this JSON snippet with the correct details for your brand product. This script tag can be added anywhere on the page.

The SEO Task Force would like to ensure that capacity to enable Rich Snippets is easily configurable. In addition, all copy that is relevant to a market and presented within the Rich Snippets should be easily translated by the business teams and/or their maintenance agencies. This rule directly applies to the new platform pilot build, Bobbi Brown Mexico.

3.8 Metadata

Nearly all metadata is invisible to your visitors and customers. Metadata is information about the content you have on your site. The SEO Task Force requires that every page holds the relevant metadata and that the data is easily updated by any business users with the right permissions.

3.8.1 Meta Description

The description attribute for the meta tag describes the page content in a summary. Best practice dictates that descriptions need to be at least 11 words but no longer than 150 characters, ideally including the page's meta keywords. Meta description appears on the SERP and is customer-facing, so should be crafted with this in mind.

In general, the business is weary of duplicate meta descriptions. Some search admin consoles will outline duplicate descriptions as an issue.

3.8.2 Meta Keywords

Meta keywords ceased to impact off site search rankings in 2009. Bing may still use them, but only as a spam signal. In other words, too many irrelevant keywords (keyword stuffing) in this attribute may actually harm rankings in Bing. However, on-site search does reference meta keywords within the internal search algorithm. And therefore should not be removed altogether. This should be considered when SEO architecture.

3.8.3 Page Titles /Title Tags

The best title tags begin with the most relevant keywords, product name, or content name, and end with the name of the site. Best practice dictates that page titles stay within 60 to 65 characters and keep the most unique, relevant, and valuable keywords toward the beginning of the title for maximum SEO benefit. Page titles appear on the SERP and are customer-facing, so should be crafted with this in mind.

3.9 Image Alts, SEO & Accessibility (PHASE 2, EXECUTION HAS NOT BEEN COMPLETED ON US SITES)

Search engines, like Googlebot, do not see images directly. They generally concentrate on the information provided in the "alt" attribute. We want to ensure that only relevant informative images have their alt text outlined as a substitute

for non-text content. SEO & accessibility guidelines share a lot of similar rules as it pertains to website usability. Screen readers for accessibility will skip all null value image alt attributes. In general, any image that is displayed with a purpose *should* have image alternative text listed with it.

If you have images in your design that are *purely* there for design reasons or the message already exists with readable copy; reference that image with CSS or make sure the HTML reference for that image is coded with an empty alt attribute. All image tags **must** include the alt attribute. The alt attribute value can be empty as not all images will need to have a text alternative. We need to balance the experience of our disabled consumer with our general business needs. The consequences of both actions will need to be analyzed.

Questions to help decide if an image should contain alt text:

- Is this image informative or decorative?
- Is this decorative image evergreen or temporary?
- Do I want this image to be searchable?
- Does this image convey a message that is not described elsewhere on the page?
- Is this image meant to be visually compelling? Are the content and message already shared in surrounding text?
- Can I add the alt copy at a later time, if required?

3.10 XML Sitemap

Sitemaps enable search engines to discover all pages on a site and to download them quickly when they change. This functionality will be available via Drupal as of September 2017.

3.10.1 Sitemap Rules

The XML includes the following Drupal pages based on their content type:

- Page Content Type
- Collection Content Type
- MPP Content Type
- Also it pulls in any page viewed as a single page with a vanity URL as well as any product URLs (either .NET generated or Drupal).

The following pages are not included in the XML sitemap and have been excluded because of SEO best practice:

- On-site search URLs
- Checkout URLs
- Account URLs
- Canonical Product URLs
- Following Directories:
 - /shared
 - /includes
 - /themes
 - /misc
 - /admin
 - /user
 - /scripts

Example DrupalGem sitemap: <http://e.maccosmetics.uk.uk.cms.elcdev.net/sitemap.xml>

3.10.2 HREFLANG & SEO

What is the hreflang tag meant for?

Hreflang is a method to mark up pages that are similar in meaning but aimed at different languages and/or regions. You can use this for three types of variations:

- Content with regional variations like en-us and en-gb.
- Content in different languages like en, de and fr.
- A combination of different languages and regional variations.

You can use hreflang to target different markets that use the same language. Using hreflang you can differentiate between the US and the UK, or between Germany and Austria.

How can it help international SEO?

If there is a version of a page that has been optimized for the users language and location, you want them to land on that page. This will help drive the local relevance of a branded site within a market. Hreflang also helps prevent the duplicate content problem. You might have the same content in English on different URLs aimed at the UK, the US and Australia. The difference on these pages might be as small as a change in prices and currency. Google might not understand on its own what you're trying to do and see it as duplicate content. With hreflang you make it very clear to the search engine that it's (almost) the same content, just optimized for different people.

Important Engineering Note

One thing is very important when implementing hreflang: don't be too specific. Let's say there are three types of pages:

- German
- German, specifically aimed at Austria
- German, specifically aimed at Switzerland

You could choose to add three hreflang attributes like this:

- de-de targeting German speakers in Germany
- de-at targeting German speakers in Austria
- de-ch targeting German speakers in Switzerland

However, which of these three results should Google show to someone searching in German in Belgium?

The first page would probably be the best. To make sure that every German searching user who does not match either de-at or de-ch gets that one, change that hreflang attribute to just de. Specifying *just* the language is in many cases a the better solution.

When you create sets of links like this, the most specific one wins. The order in which the search engines sees the links doesn't matter, it'll always try to match from most specific to least specific.

The International Request

To improve brand site market SEO rankings a "hreflang" tag should be added to our websites via the **XML sitemap** to redirect traffic to the appropriate language/market site. Currently, the only DrupalGem sites that have this existing are the EMEA Hub sites.